WHAT IS CLAIMED IS:

1.   A method for identifying and retaining customers who are likely to terminate an existing loan contract at a date earlier than the loan termination date identified in the loan contract, said method comprising the steps of:

using an early termination model to identify customers likely to terminate an existing loan contract at a date earlier than the loan termination date identified in the loan contract;

notifying those customers of new loan opportunities at competitive loan interest rates; and

cross selling new loans to the customers before the customer early terminates their existing loan.

2.   A method according to Claim 1 wherein said step of using an early termination model to identify customers likely to terminate an existing loan contract further comprises the step of utilizing six months of performance and demographic data to predict likely loan terminations at least three months before the customer is likely to prepay the existing loan.

3.   A method according to Claim 1 wherein said step of using an early termination model to identify customers likely to terminate an existing loan contract further comprises the steps of:

predicting a probability of early termination at every stage of the lifetime of the existing loan;

combining the probability of early termination with financial calculations to determine an expected profit at each stage of the existing loan;

comparing expected profit from the existing loan with an expected profit from a new loan; and

determining which customers to target for a new loan based upon the expected profits from the existing and the new loans.

4.   A method according to Claim 3 wherein said step of predicting a probability of early termination further comprises the steps of predicting probabilities using probability properties $P(A_1 \cup A_2 \cup A_3) =$

$1 - P\left((A_1 \cup A_2 \cup A_3)^C \right) = 1 - P(A_1^C \cap A_2^C \cap A_3^C) =$

$1 - P(A_1^C) P(A_2^C \mid A_1^C) P(A_3^C \mid A_1^C \cap A_2^C)$, where the events $A_1^C, A_2^C \mid A_1^C, A_3^C \mid A_1^C \cap A_2^C$ represent the probabilities that the customer has not early terminated.

5.   A method according to Claim 4 wherein said step of combining the probability of early termination with financial calculations further comprises the step of calculating an expected income, given a set of probabilities according to

$$E(Income) = (1 - P_1)I_1 + (1 - P_1)(1 - P_2)I_2 + \ldots + (1 - P_1)\ldots(1 - P_T)I_T =$$

$$= \sum_{i=1}^{T} \left[\prod_{k=1}^{i} (1 - P_k) \right] * I_i \; .$$

6.   A method according to Claim 5 wherein said step of comparing expected profit from the existing loan with an expected profit from a new loan comprises the step of calculating a monthly internal rate of return for each customer according to:

$$\sum_{t=1}^{T} \frac{C_t}{(1 + I_{RR})^t} = 0 \; .$$

7.   A method according to Claim 1 wherein said step of using an early termination model to identify customers likely to terminate further comprises the step of compiling customer data from multiple sources including at least one of world wide web, legacy customer data, facsimile, telephone and cellular data.

8.   A method according to Claim 1 wherein said step of cross selling new loans to the customers further comprises the step of cross selling at least

one of insurance product, loan products different from existing loan product and refinancing of existing loan product.

9.   A method according to Claim 1 further comprising the step of building logistic regression model to test accuracy and a strength of association between variables which describe customers and their loans within the early termination model for an ith experimental unit according to:

$$\Pi_i = F(\beta_0 + \sum_{j=1}^{J} \beta_j z_{ji}) = \frac{\exp\left(\beta_0 + \sum_{j=1}^{J} \beta_j z_{ji}\right)}{1 + \exp\left(\beta_0 + \sum_{j=1}^{J} \beta_j z_{ji}\right)}$$

where $F(\beta_0 + \sum_{j=1}^{J} \beta_j z_{ji})$ stands for the cumulative density function for a logistic distribution.

10.   A method according to Claim 9 further comprising the step of evaluating associations between variables within a matrix based on concordance according to

$$C = \sum_{i=1}^{I} \sum_{j=1}^{J} X_{ij} \left[\sum_{k>i} \sum_{l>j} X_{kl} + \sum_{k<i} \sum_{l<j} X_{kl} \right]$$

where X_{ij} stands for the number of observations for i^{th} row and j^{th} column, or disconcordance, when the subject rankings are higher on the row variable and lower on the column variable according to

$$D = \sum_{i=1}^{I} \sum_{j=1}^{J} X_{ij} \left[\sum_{k>i} \sum_{l<j} X_{kl} + \sum_{k<i} \sum_{l>j} X_{kl} \right]$$

or when the subject rankings are tied on the row variable and the column variable according to

$$T = \sum_{i=1}^{I} X_{ii} \left[\sum_{k=j \neq i} X_{kk} \right].$$

11. A method according to Claim 10 further comprising the step of ensuring that parameter estimates β_k in the logistic regression model have comparable magnitudes for the different independent variables x_k, which have different units by standardizing all continuous variables according to:

$$\widetilde{X} = \frac{X - E(X)}{\sigma(X)}$$

where $E(X)$ is the mean of X and $\sigma(X)$ is the standard deviation of X.

12. A system configured for identifying and retaining customers who are likely to terminate an existing loan contract at a date earlier than the loan termination date identified in the loan contract, said system comprising:

a server;

a network; and

at least one computer connected to said server via said network, said server configured to utilize an early termination model to identify customers likely to terminate an existing loan contract at a date earlier than the loan termination date identified in the loan contract, notify those customers of new loan opportunities at competitive loan interest rates and attempt to cross sell new loans to the customers before the customer early terminates their existing loan.

13. A system according to Claim 12 wherein said server configured to utilize loan performance and demographic data to predict likely loan terminations at least three months before the customer is likely to prepay the existing loan.

14. A system according to Claim 12 wherein said server configured to:

predict a probability of early termination at every stage of the lifetime of the existing loan;

combine a probability of early termination with financial calculations to determine an expected profit at each stage of the existing loan;

compare expected profit from the existing loan with an expected profit from a new loan; and

determine which customers to target for a new loan based upon the expected profits from the existing and the new loans.

15. A system according to Claim 12 wherein said server is configured to predict a probability of early termination using probability properties

$$P(A_1 \cup A_2 \cup A_3) = 1 - P\left((A_1 \cup A_2 \cup A_3)^C \right) = 1 - P(A_1^C \cap A_2^C \cap A_3^C) =$$
$$1 - P(A_1^C)P(A_2^C \mid A_1^C)P(A_3^C \mid A_1^C \cap A_2^C),$$ where the events $A_1^C, A_2^C \mid A_1^C, A_3^C \mid A_1^C \cap A_2^C$

represent the probabilities that the customer has not early terminated.

16. A system according to Claim 15 wherein said server is configured to calculate an expected income, given a set of probabilities according to:

$$E(Income) = (1 - P_1)I_1 + (1 - P_1)(1 - P_2)I_2 + ... + (1 - P_1)...(1 - P_T)I_T =$$
$$= \sum_{i=1}^{T} \left[\prod_{k=1}^{i} (1 - P_k) \right] * I_i .$$

17. A system according to Claim 16 wherein said server is configured to calculate a monthly internal rate of return for each customer according to:

$$\sum_{t=1}^{T} \frac{C_t}{(1 + I_{RR})^t} = 0 .$$

18.   A system according to Claim 12 wherein said server configured to compile customer data from multiple sources including at least one of world wide web, legacy customer data, facsimile, telephone and cellular data.

19.   A system according to Claim 12 wherein said server configured to build a regression model to test accuracy and a strength of association between variables which describe customers and their loans for an ith experimental unit according to:

$$\Pi_i = F(\beta_0 + \sum_{j=1}^{J} \beta_j z_{ji}) = \frac{\exp\left(\beta_0 + \sum_{j=1}^{J} \beta_j z_{ji}\right)}{1 + \exp\left(\beta_0 + \sum_{j=1}^{J} \beta_j z_{ji}\right)}$$

where $F(\beta_0 + \sum_{j=1}^{J} \beta_j z_{ji})$ stands for the cumulative density function for a logistic distribution.

20.   A system according to Claim 19 wherein said server configured to evaluate associations between variables within a matrix based on concordance according to

$$C = \sum_{i=1}^{I} \sum_{j=1}^{J} X_{ij} \left[\sum_{k>i} \sum_{l>j} X_{kl} + \sum_{k<i} \sum_{l<j} X_{kl} \right]$$

where X_{ij} stands for the number of observations for i[th] row and j[th] column, or disconcordance, when the subject rankings are higher on the row variable and lower on the column variable according to

$$D = \sum_{i=1}^{I} \sum_{j=1}^{J} X_{ij} \left[\sum_{k>i} \sum_{l<j} X_{kl} + \sum_{k<i} \sum_{l>j} X_{kl} \right]$$

or when the subject rankings are tied on the row variable and the column variable according to

-33-

$$T = \sum_{i=1}^{I} X_{ii} \left[\sum_{k=j \neq i} X_{kk} \right].$$

21.   A system according to Claim 20 wherein said server configured to ensure that parameter estimates β_k in the logistic regression model have comparable magnitudes for the different independent variables x_k, which have different units by standardizing all continuous variables according to

$$\widetilde{X} = \frac{X - E(X)}{\sigma(X)}$$

where $E(X)$ is the mean of X and $\sigma(X)$ is the standard deviation of X.

22.   A computer programmed to:

prompt a user to start an analysis of available customer information;

predict customers likely to terminate an existing loan contract at a date earlier than the loan termination date identified in the loan contract using an early termination model; and

prompt a user to notify those customers likely to early terminate of new loan opportunities.

23.   A computer according to Claim 22 programmed to utilize loan performance and demographic data to predict which loan customers are likely to prepay an existing loan at least three months before the predicted loan prepayment.

24.   A computer according to Claim 22 programmed to:

prompt a user with a prediction of a probability of early termination at every stage of the lifetime of the existing loan;

combine a probability of early termination with financial calculations to determine an expected profit at each stage of the existing loan;

compare expected profit from the existing loan with an expected profit from a new loan; and

prompt a user with a determination of customers to target for new loans based upon calculations of expected profits from the existing and the new loans.

25. A computer according to Claim 22 programmed to predict a probability of early termination using probability properties $P(A_1 \cup A_2 \cup A_3) =$
$1 - P\left((A_1 \cup A_2 \cup A_3)^C \right) = 1 - P(A_1^C \cap A_2^C \cap A_3^C) =$
$1 - P(A_1^C)P(A_2^C \mid A_1^C)P(A_3^C \mid A_1^C \cap A_2^C)$, where the events $A_1^C, A_2^C \mid A_1^C, A_3^C \mid A_1^C \cap A_2^C$ represent the probabilities that the customer has not prepaid the loan.

26. A computer according to Claim 25 programmed to calculate an expected income, given a set of probabilities according to

$$E(Income) = (1 - P_1)I_1 + (1 - P_1)(1 - P_2)I_2 + \ldots + (1 - P_1)\ldots(1 - P_T)I_T =$$
$$= \sum_{i=1}^{T} \left[\prod_{k=1}^{i} (1 - P_k) \right] * I_i \ .$$

27. A computer according to Claim 26 programmed to calculate a monthly internal rate of return for each customer according to:

$$\sum_{t=1}^{T} \frac{C_t}{(1 + I_{RR})^t} = 0 \ .$$

28. A computer according to Claim 22 programmed to test accuracy and strength of association between customers and their loans by building a regression model for an ith experimental unit according to:

$$\Pi_i = F(\beta_0 + \sum_{j=1}^{J} \beta_j z_{ji}) = \frac{\exp\left(\beta_0 + \sum_{j=1}^{J} \beta_j z_{ji}\right)}{1 + \exp\left(\beta_0 + \sum_{j=1}^{J} \beta_j z_{ji}\right)}$$

where $F(\beta_0 + \sum_{j=1}^{J} \beta_j z_{ji})$ stands for the cumulative density function for a logistic distribution.

29.   A computer according to Claim 28 programmed to evaluate associations between variables within a matrix based on concordance according to

$$C = \sum_{i=1}^{I} \sum_{j=1}^{J} X_{ij} \left[\sum_{k>i} \sum_{l>j} X_{kl} + \sum_{k<i} \sum_{l<j} X_{kl} \right]$$

where X_{ij} stands for the number of observations for i^{th} row and j^{th} column, or disconcordance, when the subject rankings are higher on the row variable and lower on the column variable according to

$$D = \sum_{i=1}^{I} \sum_{j=1}^{J} X_{ij} \left[\sum_{k>i} \sum_{l<j} X_{kl} + \sum_{k<i} \sum_{l>j} X_{kl} \right]$$

or when the subject rankings are tied on the row variable and the column variable according to

$$T = \sum_{i=1}^{I} X_{ii} \left[\sum_{k=j \neq i} X_{kk} \right].$$

30.   A computer according to Claim 29 programmed to ensure that parameter estimates β_k in the logistic regression model have comparable magnitudes for the different independent variables x_k, which have different units by standardizing all continuous variables according to

$$\widetilde{X} = \frac{X - E(X)}{\sigma(X)}$$

where $E(X)$ is the mean of X and $\sigma(X)$ is the standard deviation of X.

31. A computer readable medium comprising:

a record of customer data;

a plurality of rules for using the customer data to identify which customers are likely to early terminate a loan contract; and

a record of customers identified from said plurality of rules.

32. A computer readable medium according to Claim 31 wherein said plurality of rules identify customers likely to early terminate at least three months before the customer is likely to terminate the loan contract.

33. A computer readable medium according to Claim 31 wherein said plurality of rules comprises:

rules to predict a probability of early termination at every stage of an existing loan;

rules to combine the probability of early termination with financial calculations to determine an expected profit at every stage of the loan; and

rules to compare expected profit from the existing loan to an expected profit from a new loan.

34. A computer readable medium according to Claim 33 further comprising rules to identify which customers to target for a new loan based upon expected profits form the existing and the new loans.

35. A computer readable medium according to Claim 31 wherein said plurality of rules for using the customer data further comprises rules to cluster groups of customers according to customer characteristics.

36. A computer readable medium according to Claim 31 wherein said plurality of rules for using the customer data further comprises rules to measure a strength of association between variables which describe customers and their loans.

37. A method for identifying customers likely to terminate an existing loan contract at a date earlier than the loan termination date identified in the loan contract, said method comprising the steps of:

selecting, from a electronic interface, a record of customer information;

selecting, from the electronic interface, an early termination model for an analysis of the selected customer information; and

requesting, from the electronic interface, a list of customers identified by the model as being likely to early terminate their existing loans.

38. A method according to Claim 37 further comprising the step of requesting, from the electronic interface, that the identified customers be notified of new loan opportunities.

39. A method according to Claim 37 further comprising the step of requesting, from the electronic interface, which of the identified customers be notified of new loan opportunities.

40. A method according to Claim 37 wherein said step of selecting a record of customer information further comprises the step of selecting a source of customer information from at least one of world wide web, legacy customer data, facsimile, telephone and cellular data.

41. Apparatus comprising:

means for storing customer information;

-38-

means for identifying customers likely to terminate an existing loan contract at a date earlier than the loan termination date identified in the loan contract; and

means for prompting a user with an identification of those customers likely to early terminate.

42. Apparatus according to Claim 41 further comprising means for notifying customers likely to early terminate of new loan opportunities.

43. Apparatus according to Claim 41 wherein said means for identifying customers likely to terminate an existing loan contract comprises means for using loan performance and demographic data to identify customers likely to early terminate at least three months before the customer is likely to terminate the existing loan.

44. Apparatus according to Claim 41 wherein said means for identifying customers likely to terminate an existing loan contract comprises:

means for predicting a probability of early termination at every stage of the lifetime of the existing loan;

means for combining the probability of early termination with financial calculations to determine an expected profit at each stage of the existing loan;

means for comparing expected profit from the existing loan with an expected profit from a new loan; and

means for determining which customers to target for a new loan based upon the expected profits from the existing and the new loans.

45. Apparatus according to Claim 41 wherein said means for identifying customers likely to early terminate comprises means for using a cluster analysis to segment customers and identify sales opportunities for both active and inactive customers.

46. Apparatus according to Claim 41 wherein said means for identifying customers likely to early terminate comprises means for measuring a strength of association between variables which describe customers and their loans.

47. Apparatus according to Claim 41 wherein said means for identifying customers likely to early terminate comprises means for clustering groups of customers according to characteristics of the customers.

48. A database comprising:

data corresponding to customer information;

data corresponding to rules for using the customer information to identify which customers are likely to early terminate a loan contract; and

data corresponding to customers identified from said data corresponding to the rules.

49. A database according to Claim 48 wherein said data corresponding to rules for using customer information comprises data corresponding to rules for identifying customers likely to early terminate at least three months before the customer is likely to terminate the loan contract.

50. A database according to Claim 48 wherein said data corresponding to rules for using customer information comprises:

data corresponding to rules for predicting a probability of early termination at every stage of an existing loan;

data corresponding to rules for combining the probability of early termination with financial calculations to determine an expected profit at every stage of the loan; and

data corresponding to rules for comparing expected profit from the existing loan to an expected profit from a new loan.

51. A database according to Claim 50 further comprising data corresponding to rules for identifying which customers to target for a new loan based upon expected profits form the existing and the new loans.

52. A database according to Claim 48 wherein said data corresponding to rules for using the customer information further comprises data corresponding to rules for clustering groups of customers according to customer characteristics.

53. A database according to Claim 48 wherein said data corresponding to rules for using the customer information further comprises data corresponding to rules for measuring a strength of association between variables which describe customers and their loans.

54. A database comprising:

data corresponding to customer information; and

at least one rule to be applied to the customer information to identify customers likely to early terminate a loan.

55. A database according to Claim 54 wherein said rule to identify customers likely to early terminate a loan comprises predicting probabilities using probability properties $P(A_1 \cup A_2 \cup A_3) =$

$1 - P\left((A_1 \cup A_2 \cup A_3)^C \right) = 1 - P(A_1^C \cap A_2^C \cap A_3^C) =$

$1 - P(A_1^C)P(A_2^C \mid A_1^C)P(A_3^C \mid A_1^C \cap A_2^C)$, where the events $A_1^C, A_2^C \mid A_1^C, A_3^C \mid A_1^C \cap A_2^C$ represent the probabilities that the customer has not early terminated.

56. A database according to Claim 55 wherein said rule to identify customers likely to early terminate a loan comprises calculating an expected income for a loan, given a set of probabilities according to

$$E(Income) = (1-P_1)I_1 + (1-P_1)(1-P_2)I_2 + ... + (1-P_1)...(1-P_T)I_T =$$

$$= \sum_{i=1}^{T} [\prod_{k=1}^{i}(1-P_k)]*I_i \quad .$$

57.   A database according to Claim 56 wherein said rule to identify customers likely to early terminate a loan comprises comparing expected profit from the existing loan with an expected profit from a new loan using a calculated monthly internal rate of return for each customer calculated according to:

$$\sum_{t=1}^{T} \frac{C_t}{(1 + I_{RR})^t} = 0 \quad .$$